

OM nucleic - nucleic search, using sw model

Run on: November 22, 2001, 20:20:38 : Search time 177.69 Seconds
 (wtihout alignments)
 88411.296 Million cell updates/sec

Title: US-09-578-827A-4

Perfect score: 25/2

Sequence: 1 Aagaagcagagcgatgggttt.....ttcattaaaaaaaatg 25/2

scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0

searched: 73101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first: 45 summaries

Database :

N_Geneseq_0601:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1980.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1981.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1982.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1983.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1984.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1985.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1986.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1987.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1988.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1989.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1990.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1991.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1992.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1993.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1994.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1995.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1996.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1997.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1998.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA1999.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA2000.DAT:*, /SIDS1/gcdata/geneseq/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	203.4	8.1	936 22 AAF58252	Oligonucleotide D1
2	203.4	8.1	936 22 AAF58254	Oligonucleotide D1
3	203.4	8.1	936 22 AAF58257	Oligonucleotide D1
4	203.4	8.1	936 22 AAF58259	Oligonucleotide D2
5	203.4	8.1	936 22 AAF58262	Oligonucleotide D2
6	203.4	8.1	938 22 AAF58255	Oligonucleotide D1
7	195.4	7.8	936 22 AAF58256	Oligonucleotide D1
8	195.4	7.8	936 22 AAF58254	Oligonucleotide D1
9	195.4	7.8	936 22 AAF58257	Oligonucleotide D1
10	195.4	7.8	936 22 AAF58259	Oligonucleotide D2
11	195.4	7.8	936 22 AAF58262	Oligonucleotide D2

RESULT : 1

ID AAF58252
 XX
 AC AAF58252;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Oligonucleotide D1835.
 XX
 KW Electron-transfer group; ETM; mismatch; genotyping;
 KW gene expression; ss.
 XX
 OS Synthetic.
 XX
 PN N0200107665-A2.
 XX
 PD 01-FEB-2001.
 XX
 PP 26-JUL-2000; 2000WO-US20476.
 XX
 PR 26-JUL-1999; 99US-0145695.
 XX
 PR 17-MAR-2000; 2000US-0190259.
 XX
 PA (CLIN-) CLINICAL MICRO SENSORS INC.
 XX
 PI Umek RM;
 XX
 DR WPI; 2001-1597/28/16.
 XX
 PT Nucleic acids containing electron-transfer group, useful as labels in
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
 PT a single surface
 XX

Page 9

xx
cc
cc
cc
cc
cc
cc
xx
sq
Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other:
The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.

CC gene. In this invention an apoptosis-resistant 293 cell line (having an CC apoptosis resistant gene introduced) is established and overcomes the CC problem. The present sequence represents the base sequence of the plasmid pRx-Ires-bsr, which contains the cowpox virus bsr gene, and CC is used in an example from the present invention.

XX Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 other;

Query	Match	Score	DB	Length
Qy	496 stratacctaapccagragacataatggataaatggacacacacacaaaaggaa	2.3%	20	6644
Db	4041 aaa	46.0%	0	0
Qy	556 atcttaataaagtgttataagaataataatgccttcagggtggatccaa			
Db	4101 aaa			
Qy	616 aagggttatatgactcttaacaaaaatggcatgtacttagttggggat			
Db	4161 aaa			
Qy	676 acaagtagggattttggaaacaagaaacaaacaaangatgcgtccat			
Db	4221 aaa			
Qy	736 ttacacatgtgtttataataggtaagaaatatgacacatgggttgtt			
Db	4281 aaa			
Qy	796 ttaatgtgaagaaaaaatatgtatatttgagaatcttaaatatgtt			
Db	4341 aa			
Qy	856 ttt			
Db	4401 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaatgtgtcgaaatgttcc			
Qy	916 caaagcca 923			
Db	4461 ctaaggaa 4468			

Search completed: November 22, 2001, 22:03:13
Job time: 6555 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2001, 13:44:08 ; Search time 92.89 Seconds

(without alignments) 6100.204 Million cell updates/sec

Title: US-09-578-827A-4
perfect score: 2502
Sequence: 1 agaaagagagcggtggggtt.....ttcattaaaaaaaaatg 2502

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA,*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCNUC_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

Result No.	Score	Query Match Length	DB ID	Description
C 1	96.2	3.8	7218 1	US-08-232-463-14
C 2	96.2	2.2	19124 2	US-08-487-826B-13
C 3	49.8	2.0	665 2	US-08-488-95A-36
C 4	48.8	2.0	19124 2	US-08-487-826B-13
C 5	47.2	1.9	7218 1	US-08-432-433-14
C 6	47.2	1.9	8920 2	US-08-446-835A-1
C 7	47.2	1.9	8920 4	US-09-150-741-1
C 8	46.8	1.9	289 4	US-09-007-005-17
C 9	46.8	1.9	289 4	US-09-244-96-17
C 10	46.8	1.9	9048 3	US-09-973-273-4
C 11	4.4	1.8	827 4	US-08-998-416-535
C 12	43.6	1.7	782 4	US-08-998-416-224
C 13	43.4	1.7	5655 2	US-08-989-478-1
C 14	43.4	1.7	5655 3	US-08-996-635-1
C 15	43.4	1.7	5655 3	US-08-880-479-2
C 16	43.4	1.7	6243 2	US-09-056-075-1
C 17	43.4	1.7	6768 1	US-08-107-755A-1
C 18	43.4	1.7	8457 1	US-07-991-867B-1
C 19	43.4	1.7	8457 2	US-08-544-332-1
C 20	43.4	1.7	9919 3	US-08-880-179-1
C 21	43.4	1.7	7244 4	US-08-313-313-26
C 22	42.2	1.7	2110 4	US-09-119-159-1
C 23	41.4	1.7	2454 1	US-08-359-606-3
C 24	41.4	1.7	4507 2	US-08-459-459A-3
C 25	41.4	1.7	4507 2	US-08-887-826B-3
C 26	41.4	1.7	6050 5	RCT-US960984307
C 27	41.4	1.7	6152 4	US-08-973-462-1

RESULT 1
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:	
APPLICANT:	DORNER, F.
APPLICANT:	SCHEIFFLINGER, F.
APPLICANT:	FALKNER, F. G.
TITLE OF INVENTION:	RECOMBINANT FOWPOX VIRUS
NUMBER OF SEQUENCES:	52
CORRESPONDENCE ADDRESS:	
ADDRESSEE:	Foley & Lardner
STREET:	1800 Diagonal Road, Suite 500
CITY:	Alexandria
STATE:	VA
COUNTRY:	USA
ZIP:	22313-0299

COMPUTER READABLE FORM:	
COMPUTER TYPE:	Floppy disk
OPERATING SYSTEM:	PC-DOS/MS-DOS
SOFTWARE:	PatentIn Release #1.0, version #1.25
CURRENT APPLICATION DATA:	APPLICATION NUMBER: US/08/232,463
FILING DATE:	CLASSIFICATION: 435
PRIOR APPLICATION DATA:	APPLICATION NUMBER: US/07/935,313
FILING DATE:	APPLICATION NUMBER: EP 91 114 300.6
FILING DATE:	ATTORNEY/AGENT INFORMATION:
NAME:	BENT, Stephen A.
REGISTRATION NUMBER:	29,768
REFERENCE DOCKET NUMBER:	30472/114 IMMU
TELECOMMUNICATION INFORMATION:	TELEPHONE: (703) 836-9300
TELEFAX:	(703) 603-4109
TELEX:	899149
INFORMATION FOR SEQ ID NO: 14:	SEQUENCE CHARACTERISTICS: LENGTH: 7218 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	single
TOPOLOGY:	linear
IMMEDIATE SOURCE:	CLONE: ptzgpt-FLS
SEQUENCE 1:	US-08-232-463-14

Query Match ; TELEFAX: (619) 235-0176
 Best Local Similarity 3.8%; Score 96.2; DB 1; Length 7218;
 Matches 35; Conservative 254; Mismatches 152; Indels 0; Gaps 0;

SEQUENCE CHARACTERISTICS:
 LENGTH: 19124 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-487-826B-13

Query Match ; INFORMATION FOR SEQ ID NO: 13:
 Best Local Similarity 2.2%; Score 56; DB 2; Length 19124;
 Matches 304; Conservative 0; Mismatches 370; Indels 6; Gaps 2;

GENERAL INFORMATION:
 APPLICANT: Sim, Kim L.
 APPLICANT: Chittnis, Chetan
 APPLICANT: Miller, Louis H.
 APPLICANT: Peterson, David S.
 APPLICANT: Su, Xin-zhuan
 APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US08/487,826B
 FILING DATE: 10-SEP-1993
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Israelien, Ned
 REGISTRATION NUMBER: 29, 655
 REFERENCE/DOCKET NUMBER: NIH121.001CP1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550

RESULT ;
 US-08-487-826B-13/c Sequence 13, Application US/08487826B
 Patent No. 593827

GENERAL INFORMATION:
 APPLICANT: Sim, Kim L.
 APPLICANT: Chittnis, Chetan
 APPLICANT: Miller, Louis H.
 APPLICANT: Peterson, David S.
 APPLICANT: Su, Xin-zhuan
 APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US08/487,826B
 FILING DATE: 10-SEP-1993
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Israelien, Ned
 REGISTRATION NUMBER: 29, 655
 REFERENCE/DOCKET NUMBER: NIH121.001CP1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550

RESULT ;
 US-08-487-826B-13/c Sequence 36, Application US/08883795A
 Patent No. 5985607

GENERAL INFORMATION:
 APPLICANT: Delcive, Genevieve
 APPLICANT: Awang, Gregor

RESULT 14
US-08-996685-1/c
Sequence 1, Application US/08896685
GENERAL INFORMATION:
 Patent No. 6031153
APPLICANT: Ryals, John
 Friedrich, Leslie
APPLICANT: UKnes, Scott
APPLICANT: Molina, Antonio
APPLICANT: Russ, Wilhelm
APPLICANT: Knorr-Batler, Gertrude
APPLICANT: Kung, Ruth
APPLICANT: Kessmann, Helmut
APPLICANT: Oostendorp, Michael
TITLE OF INVENTION: METHOD FOR PROTECTING PLANTS
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 603115artis Corporation
 STREET: 3054 Cornwallis Road
 CITY: Research Triangle Park
 STATE: No. 6031153th Carolina
 COUNTRY: USA
 ZIP: 27709
COMPUTER READABLE FORM:
 COMPUTER TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, version #1.30
CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/996,685
 APPLICATION NUMBER: US/08/996,685
 FILING DATE:
CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/761,543
 FILING DATE: 6-DEC-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/034,378
 FILING DATE: 27-DEC-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/034,379
 FILING DATE: 27-DEC-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/034,382
 FILING DATE: 27-DEC-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/034,730
 FILING DATE: 10-JAN-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/035,021
 FILING DATE: 10-JAN-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/035,022
 FILING DATE: 10-JAN-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/035,024
 FILING DATE: 10-JAN-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/875,015
 FILING DATE: 16-JUL-1997
ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: PF/5-21215/P1/CGC1912
TELECOMMUNICATION INFORMATION:
 TELEPHONE: (919) 541-8587
 TELEFAX: (919) 541-8689
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5655 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
FEATURE:
 NAME/KEY: exon
 LOCATION: 2787..3347
 OTHER INFORMATION: /Product= "1st exon of NIM1"
FEATURE:
 NAME/KEY: exon
 LOCATION: 4271..4474
 OTHER INFORMATION: /product= "3rd exon of NIM1"
FEATURE:
 NAME/KEY: exon
 LOCATION: 4586..4866
 OTHER INFORMATION: /product= "4th exon of NIM1"
FEATURE:
 NAME/KEY: CDS
 LOCATION: Join(2787..3347, 3427..4162, 4271..4474, 4586..4866)
RESULT 15
US-08-880-179-2/c
Sequence 2, Application US/08880179
GENERAL INFORMATION:
 Patent No. 6031004
APPLICANT: Ryals, John
APPLICANT: Delaney, Terry
APPLICANT: Friedrich, Leslie
APPLICANT: Weymann, Kristianna
APPLICANT: Lawson, Ray
APPLICANT: Ellis, Daniel
APPLICANT: UKnes, Scott
APPLICANT: Jesse, Taco
APPLICANT: Vos, Pieper
TITLE OF INVENTION: GENE ENCODING A PROTEIN INVOLVED IN THE SIGNAL TRANSDUCTION CASCADE LEADING TO SYSTEMIC ACQUIRED RESISTANCE
TITLE OF INVENTION: IN PLANTS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 6031004artis Corporation
 STREET: 520 White Plains Road, P.O. Box 2005
 CITY: Tarrytown
 STATE: New York
 COUNTRY: USA
 ZIP: 10591
COMPUTER READABLE FORM:

LOCATION: JOIN(4/6/.334/, 342/.410/, 42/1.44/, 4300.4000)

Search completed: November 22, 2001, 22:01:01
Job time: 29813 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2001, 13:20:33 ; Search time 1858.82 Seconds
(without alignments)
12723.685 Million cell updates/sec

Title: US-09-578-827A-4
Perfect score: 2502
Sequence: 1 agaaggcagcggtgggttt.....ttcataaaaaaaatg 2502

Scoring table: IDENTITY.NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : EST+
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estrol1:*
83: em_estrol2:*
84: em_estrol3:*
85: em_estrol4:*
86: em_estrol5:*
87: em_estrol6:*
88: em_estrol7:*
89: em_estrol8:*
90: em_estrol9:*
91: em_estrol0:*
92: em_estrol1:*
93: em_estrol2:*
94: em_estrol3:*
95: em_estrol4:*
96: em_estrol5:*
97: em_estrol6:*
98: em_estrol7:*
99: em_estrol8:*
100: em_estrol9:*
101: em_estrol0:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

```

117: qb_est48:*
118: qb_est49:*
119: qb_est50:*
120: qb_est51:*
121: qb_est52:*
122: qb_est53:*
123: qb_est54:*
124: qb_est55:*
125: qb_est56:*
126: qb_est57:*
127: qb_est58:*
128: qb_est59:*
129: qb_est60:*
130: qb_est61:*
131: qb_est62:*
132: qb_est63:*
133: qb_est64:*
134: qb_est65:*
135: qb_est66:*
136: qb_est75:*
137: qb_est76:*
138: qb_est77:*
139: qb_est78:*
140: qb_est79:*
141: qb_est80:*
142: qb_est81:*
143: qb_est82:*
144: qb_est83:*
145: qb_est84:*
146: qb_est85:*
147: qb_est86:*
148: qb_est87:*
149: qb_est88:*
150: qb_est89:*
151: qb_est90:*
152: qb_est99:*
153: qb_est100:*
154: qb_est101:*
155: qb_est102:*
156: qb_est103:*
157: qb_est104:*
158: qb_est115:*
159: qb_est106:*
160: qb_est67:*
161: qb_est68:*
162: qb_est69:*
163: qb_est70:*
164: qb_est71:*
165: qb_est72:*
166: qb_est73:*
167: qb_est74:*
168: qb_est91:*
169: qb_est92:*
170: qb_est93:*
171: qb_est94:*
172: qb_est95:*
173: qb_est96:*
174: qb_est97:*
175: qb_est98:*
176: em_esthun30:*
177: em_esthun31:*
178: em_esthun32:*
179: em_esthun33:*
180: em_estom3:*
181: em_estom31:*
182: em_estp111:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: qb_est107:*
188: qb_est108:*
189: qb_est109:*
190: qb_est110:*
191: qb_est111:*
192: qb_htc:*
193: em_gss_fun:*
194: em_gss_num1:*
195: em_gss_num2:*
196: em_gss_num3:*
197: em_gss_num4:*
198: em_gss_hun5:*
199: em_gss_hun6:*
200: em_gss_hun7:*
201: em_gss_hun8:*
202: em_gss_hun9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pnl1:*
208: em_gss_pnl2:*
209: em_gss_pro:*
210: em_gss_rdl1:*
211: em_gss_rdl2:*
212: em_gss_rdl3:*
213: em_gss_rdl4:*
214: em_gss_rdl5:*
215: em_gss_rtl1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: qb_gss1:*
219: qb_gss2:*
220: qb_gss3:*
221: qb_gss4:*
222: qb_gss5:*
223: qb_gss6:*
224: qb_gss7:*
225: qb_gss8:*
226: qb_gss9:*
227: qb_gss10:*
228: qb_gss11:*
229: qb_gss12:*
230: qb_gss13:*
231: qb_gss14:*
232: qb_gss15:*
233: qb_gss16:*
234: qb_gss17:*
235: qb_gss18:*
236: qb_gss19:*
237: qb_gss20:*
238: qb_gss21:*
239: qb_gss22:*
240: qb_gss23:*
241: qb_gss24:*
242: qb_gss25:*
243: qb_gss26:*
244: qb_gss27:*
245: qb_gss28:*
246: qb_gss29:*
247: qb_gss30:*
248: qb_gss31:*
249: qb_gss32:*
250: qb_gss33:*
251: qb_gss34:*
252: em_gss_inv4:*
253: em_gss_rdl6:*
254: em_gss_rdl7:*
255: em_gss_rdl8:*
256: qb_gss35:*
257: qb_gss36:*
258: qb_gss37:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES
JOURNAL

Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage : BP 191 91006 Evry Cedex - FRANCE (E-mail : seqref@genoscope.cnrs.fr)

COMMENT

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.irutify.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kazutoyo Osoegawa and Aaron Mummery in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPC1-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the Isogenic strain Y2, on bw sp, the same strain used for the BDGP's P1 and ESN libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

DESCRIPTION

Result No.	Score	Query Match	Length	DB	ID	Description
SUMMARIES						
C 1	87.6	3.5	1.01	219	CNS0039G	AL063921 <i>Drosophila melanogaster</i>
C 2	87.4	3.5	997	219	CNS005TE	AL06767 <i>Drosophila</i>
C 3	84.2	3.4	1.01	219	CNS005BD	AL06091 <i>Drosophila</i>
C 4	81.8	3.3	1.01	219	CNS0059G	AL063921 <i>Drosophila</i>
C 5	81.2	3.2	1.01	219	CNS00EVL	AL06706 <i>Drosophila</i>
C 6	77.6	3.1	974	219	CNS00IT	AL075432 <i>Drosophila</i>
C 7	76.8	3.1	1001	219	CNS0054G	AL063781 <i>Drosophila</i>
C 8	76.6	3.1	1092	220	CNS020K7	AL175596 <i>Tetradon</i>
C 9	76.6	3.1	1.01	219	CNS001T2	AL07814 <i>Drosophila</i>
C 10	76.4	3.1	1.01	219	CNS0183Y	AL108856 <i>Drosophila</i>
C 11	75.4	3.0	1092	220	CNS020K7	AL175596 <i>Tetradon</i>
C 12	75.2	3.0	1043	219	CNS0115P	AL103735 <i>Drosophila</i>
C 13	75.2	3.0	1201	219	CNS0167P	AL103936 <i>Drosophila</i>
C 14	74.8	3.0	1101	219	CNS001BD	AL064091 <i>Drosophila</i>
C 15	74.2	3.0	1101	219	CNS00EVL	AL063706 <i>Drosophila</i>
C 16	73.8	2.9	928	219	CNS00DKY	AL071865 <i>Drosophila</i>
C 17	71.8	2.9	996	219	CNS001FH	AL071063 <i>Drosophila</i>
C 18	71.6	2.9	1.01	219	CNS001J	AL063938 <i>Drosophila</i>
C 19	70.8	2.8	1.01	219	CNS01FB	AL106732 <i>Drosophila</i>
C 20	70.6	2.8	987	219	CNS014P	AL103403 <i>Drosophila</i>
C 21	70.4	2.8	844	220	B10796	B10796 T26G15-Sp6
C 22	70.4	2.8	1.01	219	CNS0056Z	AL097301 <i>Drosophila</i>
C 23	70	2.8	1.01	219	CNS0100X	AL098379 <i>Drosophila</i>
C 24	69.8	2.8	734	219	CNS010MP	AL09163 <i>Drosophila</i>
C 25	69.8	2.8	1494	141	BE889859	BE889859 601681483
C 26	69.4	2.8	1101	219	CNS000B8	AL053632 <i>Drosophila</i>
C 27	69.2	2.8	638	105	AL513901	AL513901 AL513901
C 28	68.8	2.7	959	219	CNS0055	AL06806 <i>Drosophila</i>
C 29	68.8	2.7	1029	220	CNS012GM	AL174271 <i>Tetradon</i>
C 30	68.8	2.7	1996	256	B0949	B0949 T1F7-Sp6 TA
C 31	68.6	2.7	1001	219	CNS0140O	AL103554 <i>Drosophila</i>
C 32	68.4	2.7	987	219	CNS014PQ	AL104456 <i>Drosophila</i>
C 33	68.4	2.7	1101	219	CNS002FG	AL062437 <i>Drosophila</i>
C 34	68.2	2.7	987	219	CNS00418	AL065337 <i>Drosophili</i>
C 35	68.2	2.7	1101	219	CNS017V2	AL108536 <i>Drosophili</i>
C 36	68.2	2.7	1101	219	CNS0182P	AL108811 <i>Drosophili</i>
C 37	67.8	2.7	873	220	CNS024M5	AL109505 <i>Tetradon</i>
C 38	67.8	2.7	154	154	BG520365	BG520365 ps20C9-Y
C 39	67.8	2.7	1101	219	CNS014K	AL103740 <i>Drosophili</i>
C 40	67.6	2.7	816	220	CNS02KJY	AL201604 <i>Tetradon</i>
C 41	67.6	2.7	1101	219	CNS001JT	AL058307 <i>Drosophili</i>
C 42	67.2	2.7	1101	219	CNS010KX	AL059895 <i>Drosophili</i>
C 43	67	2.7	1068	219	CNS001TV	AL059846 <i>Drosophili</i>
C 44	66.6	2.7	1201	219	CNS0165X	AL063335 <i>Drosophili</i>
C 45	66.6	2.7	1443	167	BB421656	BB421656 HWM012OD.

ALIGNMENTS

RESULT	1	CNS039G/C	LOCUS	CNS039G	1101 bp	DNA	GSS	03-JUN-1999
DEFINITION								
Drosophila melanogaster genome survey sequence TBT3 end of BAC # AL039210 of RPCI-8 library from Drosophila melanogaster (fruit fly), genomic survey sequence.								
ACCESSION								
VERSION	AL03921.1	GT:494178	KEYWORDS	GSS:				
ORGANISM	fruit fly.							
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;								
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;								
Muscomorpha; Schizophora; Drosophilidae; Drosophila;								
REFERENCE	1 (bases 1 to 1101)		AUTHORS	Genoscope.				
TITLE	Direct Submission							

RESULT

Query Match 3.0%; Score 74.8; DB 219; Length 1101;
 Best Local Similarity 41.3%; Pred. No. 0.00025;
 Matches 199; Conservative 46; Mismatches 229; Indels 8; Gaps 1

PL and EST Libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila.bac.htm>.

```

source
  1. 1101
  /organism="Drosophila melanogaster"
  /db_xref="taxon:227"
  /clone_1_id="RPCI-98"
  /clone="BACR29B23"
  /note="end : T7"
  BASE COUNT    419 a    91 c    60 g    299 t    232 others
  ORIGIN

```

Query Match 3.0%; Score 74.2; DB 219; Length 1101;
 Best Local Similarity 32.4%; Pred. No. 0.00033;
 Matches 206; Conservative 105; Mismatches 324; Indels 0; Gaps 0;

Db 1086 TTTTWWTTAWTTAWATAAAWTATTTWTTATWTTAWWWNTATWATATTTWTTW 1027
QY 837 aaaaatgttaaaaaatgttggaaaaaaatggggaaaaatagtggcagsga 896

Db	Sequence	Length
1026	TAATATATWTTATAATATAWTAWAWAGATATATAWANWTTAWAWAT	967

NS00EVL/C
OCUS CNS00EVL
1101 bp DNA GSS 04-JUN-1999

LOCUS DROSOPHILA MELANOGASTER GENOME SURVEY SEQUENCE T7 END OF BAC: BACR2B23 OF RPO1-98 LIBRARY FROM DROSOPHILA MELANOGASTER (FLY), GENOMIC SURVEY SEQUENCE.
DEFINITION AL069706
VERSION 1
SOURCE GSS
LENGTH 1101 bp
ACCESSION AL069706.1
LINKS GI:49494949

ORGANISM Drosophila melanogaster
SOURCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; fruit fly.

Bukaryota; Metzoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)

Db	Qy	Hs	Sp
606	ATTAATWAWATTTATAWATATAAWWTWTAAATTWTAAWTWWTTAATTAA	547	

TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
Determination of this barcode sequence was carried out as part of a

ג'נְדָּעָה

Search completed: November 22, 2001, 21:02:52
Job time: 3739 sec

GenCore version 4.5
 copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model
 Run on: November 22, 2001, 13:28:18 ; Search time 3308.22 Seconds
 (without alignments)
 11698.227 Million cell updates/sec

Title: US-09-578-827A-4
 Perfect score: 2502
 Sequence: 1 agaaaggagagcggtgggtt.....ttcattaaaaaaaatg 2502

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba1:*

2: gb_ba2:*

3: gb_ba3:*

4: gb_in1:*

5: gb_in2:*

6: gb_in3:*

7: gb_om:*

8: gb_ovr:*

9: gb_pat1:*

10: gb_pat2:*

11: gb_ph:*

12: gb_p11:*

13: gb_p12:*

14: gb_p13:*

15: gb_p14:*

16: em_ba1:*

17: em_ba2:*

18: em_fan:*

19: em_htg_hum:*

20: em_htg_inv:*

21: em_htg_rod:*

22: em_htg_hum1:*

23: em_htg_hum2:*

24: em_htg_hum3:*

25: em_htg_hum4:*

26: em_htg_hum5:*

27: em_htg_hum7:*

28: em_htg_hum8:*

29: em_htg_inv1:*

30: em_htg_inv2:*

31: em_htg_inv3:*

32: em_htg_other:*

33: em_htg_rod:*

34: em_hum1:*

35: em_hum2:*

36: em_hum3:*

37: em_hum4:*

38: em_hum5:*

39: em_hum6:*

40: em_hum7:*

41: em_in:*

42: em_om:*

43: em_or:*

44: em_ov:*

45: em_pat:*

46: em_ph:*

47: em_pl:*

48: em_ro:*

49: em_sts:*

50: em_sy:*

51: em_un:*

52: em_vl:*

53: gb_st1:*

54: gb_st2:*

55: gb_st3:*

56: gb_sy:*

57: gb_un:*

58: gb_vl:*

59: gb_vl2:*

60: gb_hg1:*

61: gb_hg2:*

62: gb_hg3:*

63: gb_hg4:*

64: gb_hg5:*

65: gb_hg6:*

66: gb_hg7:*

67: gb_hg8:*

68: gb_hg9:*

69: gb_hg10:*

70: gb_htg11:*

71: gb_htg12:*

72: gb_htg13:*

73: gb_htg14:*

74: gb_hg15:*

75: gb_hg16:*

76: gb_htg17:*

77: gb_htg18:*

78: gb_htg19:*

79: gb_hg20:*

80: gb_hg21:*

81: gb_hg22:*

82: gb_htg23:*

83: gb_hg24:*

84: gb_hg25:*

85: gb_pr1:*

86: gb_pr2:*

87: gb_pr3:*

88: gb_pr4:*

89: gb_pr5:*

90: gb_pr6:*

91: gb_pr7:*

92: gb_pr8:*

93: gb_pr9:*

94: gb_roi:*

95: gb_ro2:*

96: gb_in1:*

97: gb_pr0:*

98: em_ba3:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2502	100%	91740	AT19FL8
2	2502	100%	196339	ATCHRIV87
3	1001	48.0	2825	AF23752 Arabidops
4	96.2	3.8	7218	166494 Sequence 14
5	80.6	3.2	4601	U11584 Drosophila
6	80.6	3.2	19517	DMU11584
7	77.1	3.1	176174	ACU07483 Homo sapi
8	76.8	3.1	165097	AL355097 Homo sapi
9	76.8	3.1	165097	CNS05TCG

FEATURES	source	1.	1. . 9140	
COMMENT				Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosome 4 can be viewed at: http://websvr.mips.biochem.mpg.de/proj/thal/ .
RESULT	1			
ATR19F18				
LOCUS	ATP19F18	91740 bp	DNA	PLN
DEFINITION	Arabidopsis thaliana DNA chromosome 4, BAC clone	F19F18 (ESSA project).	04-MAR-1999	P19F18
ACCESSION	AI035605			
VERSION	AI035605.1	GI:4468976		
KEYWORDS	thale cress, Arabidopsis thaliana			
SOURCE	Organism			
REFERENCE	1. (bases 1 to 91740)			
AUTHORS	Berw, M., Koetter, R., Hempel, S., Entian, K.-D., Bancroft, I., Brässig, M., Koetter, R., Hempel, S., Entian, K.-D., Bancroft, I., Mewes, H.W., Mayer, K.F.X. and Schueler, C.			
JOURNAL	2 (bases 1 to 91740)			
REFERENCE	EU Arabidopsis sequencing project.			
AUTHORS	Title			
JOURNAL	Submitted (03 MAR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG. E-mail: schuelle@mps.biocbem.mpg.de,mayer@mps.biocbem.mpg.de Project Coordinator: Mike Berw, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UZ Norwich, UK, E-mail: michael.berw@bbsrc.ac.uk			
COMMENT				
FEATURES	source	1.	1. . 9140	
ALIGNMENTS				
				/organism="Arabidopsis thaliana"
				/variety="Columbia"
				/db_xref="txon:3702"
				/chromosome="4"
				gene
				gene=F19F18_10"
				/join="2053..2268,2351..2551,2625..2790,3003..3409)
				CDS
				AC012492 Homo sapi
				AL47405 Anopheles
				AC007483 Homo sapi
				AF203688 Canavalia
				AC005505 Plasmodium
				AL12783 Homo sapi
				S98842 SP96-spore
				AF211124 Carlsonell
				AC010727 Homo sapi
				AC011718 Homo sapi
				AC019041 Homo sapi
				AL022655 Caenorhab
				AL022596 Caenorhab
				Z98855 Caenorhabdi
				AC022597 Caenorhab
				AC009277 Homo sapi
				AL034559 Lasmodium
				AL035100 Homo sapi
				AC083827 Homo sapi
				AX083744 Sequence
				AC087566 Mus muscu
				U37141 Drosophila
				AL355097 Homo sapi
				AL034559 Plasmodium
				AL041457 T7 end of
				AL034559 Plasmodium
				AX083744 Sequence
				AC079409 Homo sapi
				AC009524 Plasmodium
				U11884 Drosophila
				AC006970 Homo sapi
				M23149 Dictyoscelid
				AC024285 Homo sapi
				AC073136 Homo sapi
				AL355372 Homo sapi
				AC069525 Homo sapi
				AC005089 Homo sapi
				exon
				/gene="F19F18_10"
				/number=1
				intron
				/gene="F19F18_10"
				/number=2
				exon
				/gene="F19F18_10"
				/number=2
				exon
				/gene="F19F18_10"
				/number=3
				intron
				/gene="F19F18_10"
				/number=3
				exon
				/gene="F19F18_10"
				/number=4
				gene
				5135..6431
				/gene="F19F18_20"
				/join="5135..5350,5437..5637,5727..5892,6025..6431)
				CDS
				/gene="F19F18_20"
				/product="peroxidase-like protein"
				/protein_id="CA38292_1"
				/db_xref="GI:4468978"
				/translate="GI:4468978"
				Contains Peroxidases signatures [DMTALSGAHTL]
				/codon_start=1
				/product="peroxidase-like protein"
				/protein_id="CA38292_1"
				/db_xref="GI:4468978"
				/translate="GI:4468978"
				NAVOKKIQQTPTTIPATRLRFLHDFEVNGCDASVIASTWNKAKDDHNLNSLAGD
				FDTVKRKEAVDAVPNCRNKSCADLTMATRDVNLAGSPQAYELGRDGLOSSAS
				SVTGKLPKTEDNLQNLNGSNDMIALSSQHAFCTKVFLQNLKFLNQTK
				NVDPTTNDKYVTELKASCPOQIDPRVAINMDPNTPQDFNVYVQLQKGFLPSDQV
				LFTDSRSKPTDILWANGOLFNOATFISMKTLGRGYKGKGSNGNTRDGAFT
				5135..5350
				exon
				/gene="F19F18_20"
				/number=1
				intron
				5351..5435
				exon
				/gene="F19F18_20"
				/number=1
				exon
				5437..5637
				exon
				/gene="F19F18_20"
				/number=2
				intron
				5638..5725
				exon
				/gene="F19F18_20"
				/number=2
				5727..5892

RESULT_2
 ATCHRIV87
 LOCUS ATCHRIV87 196339 bp DNA chromosome 4, contig fragment No. 87.
 DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 87.
 ACCESSION AL11591
 VERSION AL11591.2 GI:7270703
 KEYWORDS thale cress.
 SOURCE_
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viriplantae; Embryophyta; Tracheophyta; Spermatophytina

REFERENCE	Magnoliophyta; eu dicots; core eudicots; Rosidae; eurosids III
AUTHORS	Brassicaceae; Arabidopsis; 1 (bases 42610 to 13618; 123423 to 19639)
JOURNAL	Rose, M., Hempel, S., Entian, K.-D., Mewes, H.W., Lemcke, K. and Mayer, K.F.X.
REFERENCE	Unpublished
2	(bases 1 to 196339)

Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemkick@mps.bioc.chem.mpg.de, mayer@mps.bioc.chem.mpg.de, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbrc.ac.uk

COMMENT

annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.bioc.chem.mpg.de/proj/thal/> this fragment has an overlap with ATCHRIV86 at the 5' end and an overlap with ATCHRIV88 at the 3' end.

FEATURES

source	Location/Qualifiers
	1..19639
/organism="Arabidopsis thaliana"	
/variety="Columbia"	
/db_xref="taxon:3702"	
/chromosome="4"	
gene	
exon	6146..7792
CDS	/gene="AT4g37210"
	6146..6474
	/gene="AT4g37210"
	6146..6474..6702..6919..7013..7225..7313..7403..7489..7792
	/note="AT4g37210"
	/note="intron number 1 is a special U12 intron
	similarity to nuclear histone-binding protein N1/N2,
	Xenopus laevis, PIR2:A25680"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAB0387..1"
	/ab_xref="GI:7270704"
	/translation="MVERSASASEASYVIOTLREPATEIAOTPLEPLNLASIAEATVSYVQ
	GTEETCNNDNNNRAADATEEVCDERKTEFELTEKGSLVKEENPDAVDC
	FSRALKRVAHGEEDAEKNCIYVCKLAKAQAEADPQGNPKKEGVQESSENG
	ESLAPSTVSGDPERQSSGGGGKDGEGEDODDDSDAOGDADEBDSLDM
	WMLKARVTDTCMFLKIDCLCSLNELEDDESSLDYKALSLNEKGAAGSATST
	VSEBIDEQIOQSENVPYIDSKASDKEVIGDLAGLADLEKKASKLNLSVH"
intron	
	6475..6701
	/gene="AT4g37210"
	/number=1
exon	6702..6919
	/gene="AT4g37210"
	/number=2
intron	6920..7012
	/gene="AT4g37210"
	/number=2
exon	7013..7225
	/gene="AT4g37210"
	/number=3
intron	7226..7312
	/gene="AT4g37210"
	/number=3
exon	7313..7403
	/gene="AT4g37210"
	/number=4
intron	7404..7488
	/gene="AT4g37210"
	/number=4
exon	7489..7792
	/gene="AT4g37210"
	/number=5
gene	
CDS	/product="cold acclimation protein homolog"
	/protein_id="CAB80380..1"
	/ab_xref="GI:7270705"
	/translation="MGEREFAMKTEENAAUNLINSMDNEFAAKKLKVKDVGMLGGVG
	FCTSVOMAATFAVYLILLDRTNMKTMMTPIFOFGGDFCKWL
	ALITYLVRLPFKEFEWLELPVALLIVVSPSLIATLRESWVGAVICLIVACILE
	HEHIKASGRFRNSFTOKNGISNTIGIVALLYPVNTIFFHF"
	8876..9051
	/gene="AT4g37220"
	/number=1
	9052..9140
	/gene="AT4g37220"
	/number=1
	9141..9248
	/gene="AT4g37220"
	/number=2
	9249..9329
	/gene="AT4g37220"
	/number=2
	9330..9400
	/gene="AT4g37220"
	/number=3
	9401..9485
	/gene="AT4g37220"
	/number=3
	9486..9739
	/gene="AT4g37220"
	/number=4
	9739..11357
	/gene="AT4g37230"
	/number=1
	10194..10398
	/gene="AT4g37230"
	/ab_xref="GI:7270706"
	contains EST gb:Z34655"
	/codon_start=1
	/product="photosystem II oxygen-evolving complex like
	protein (partial)"
	/protein_id="CAB0389..1"
	/ab_xref="GI:7270706"
	/translation="MISIVVYDIEISGPFFEDYAAVTVNNDLVGSYRSSSSWH
	RKQYFGKSRPAFTCHVYKRSDDQPGYDIAVALLARGNINELAKENKNITLVSYK
	NPESEGVIGAFESIQPSDSTDGATTPDKVKIQGIVWCQLDE"
	/gene="AT4g37230"
	/number=1
exon	11131..11357
	/gene="AT4g37230"
	/number=2
	18308..18822
	/gene="AT4g37240"
	/number=1
	18388..18822
	/gene="AT4g37240"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAB0390..1"
	/ab_xref="GI:7270707"
	/translation="MEANPKWVQVILKYPKMFICNSDDDFDDAAVATSADEEQL
	GOYFALPCLWIRQPLKAEEMALAVASSALMRGGGCRKCVEPVISDKLRMVG
	SGDDTVSGSGRKVRKGDGGSVSSRRRKCYAELSTIDE"
	21559..23955
	/gene="AT4g37250"

Upstream BAC (overlapping the T7 end) : R-74H1 (AC=AL445594)
Downstream BAC (overlapping the SP6 end) : R-363M17_PCR1
----- Summary Statistics -----
Assembly program: Phrap; version 2.0
Quality coverage: 5.98x in Q20 bases; sum-of-contigs

Overall quality chart :
Range : bases
0 : 1
1 - 9 : 58
10 - 19 : 316
20 - 29 : 429
30 - 39 : 1054
40 - 49 : 5675
50 - 59 : 7528
60 - 69 : 8998
70 - 79 : 19376
80 - 89 : 48095
90 - 99 : 73567

Percentage of bases with a quality value >= 40 : 98 %.

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1. 165097
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="R-204N11"
/clone_id="RP01-11"
/chromosome="14"
/db_xref="taxon:9606"

FEATURES

Source	BASE COUNT	ORIGIN
	49700 a 34701 c 33268 g 47427 t	1 others

RESULT 8

LOCUS	CNS05TCG	DEFINITION	LOCUS	CNS05TCG	DEFINITION
ACCESSION	AL355097	PROGRESS ***, in Ordered pieces.	ACCESSION	AL355097	PROGRESS ***, in Ordered pieces.
VERSION	4.4	GI:1122956	VERSION	4.4	GI:1122956
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.	KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.		
SOURCE		SOURCE			
ORGANISM	Homo sapiens	ORGANISM	Homo sapiens		
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo. 1 (bases 1 to 165097)	JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo. 1 (bases 1 to 165097)		
REFERENCE		REFERENCE			
COMMENT	Submitted (17-NOV-2000) Genoscope - Centre National de Sequenage : BTP 191 9106 EVRY Cedex FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) On Nov 20, 2000 this sequence version replaced gi:10944764.	COMMENT	Submitted (17-NOV-2000) Genoscope - Centre National de Sequenage : BTP 191 9106 EVRY Cedex FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) On Nov 20, 2000 this sequence version replaced gi:10944764.		
AUTHORS		AUTHORS			
TITLE	Direct Submission	TITLE	Direct Submission		
Submitted	(17-NOV-2000) Genoscope - Centre National de Sequenage : BTP 191 9106 EVRY Cedex FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) On Nov 20, 2000 this sequence version replaced gi:10944764.	Submitted	(17-NOV-2000) Genoscope - Centre National de Sequenage : BTP 191 9106 EVRY Cedex FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) On Nov 20, 2000 this sequence version replaced gi:10944764.		
Center	Genoscope / Centre National de Sequenage	Center	Genoscope / Centre National de Sequenage		
Center code	GS	Center code	GS		
Web site	http://www.genoscope.cns.fr/	Web site	http://www.genoscope.cns.fr/		
Contact	SeqRef@genoscope.cns.fr	Contact	SeqRef@genoscope.cns.fr		

IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continue. The sequence may be contaminated with foreign sequence from E.coli, Yeast, vector, phage, etc... even if efforts are made to eliminate these contaminating sequences. The following BAC sequence is oriented from the T7 to the SP6 end.

5276 .54583)
 /gene="OOCt1"
 /note="mRNA sequence presented in GenBank Accession Number AF033562."
 /codon_start=1
 /product="Ornithine carbamoyltransferase OOCt1"
 /protein_id="AA43481.1"
 /db_xref=GI:12003188"
 /translation="MAISCCYCTGSDKPLSQRRTPLPAVSURRRISI
 ITATSAFQKVQKNCFHDDDKDTLKMRLALEVKSLKGDRKFRPEKGKTSMI
 FTKPSHRTSYFETGETTLLGHAYIGPDDQMGREEDRVARLSRNDILWARV
 NHQDILDAKTVWPPINGLTDYHNCQIMADJWICRRECTKVIYVGDNIV
 HSWLIMASVIFHFVACPKFFDPDKTVERAKRGISKIBITNPKEAVKKGADVYS
 NSIVFOAENRHMHQAMVLMHLGK"
 DYNASHKTYVPPINGLTDYHNCQIMADJWICRRECTKVIYVGDNIV
 AHDQDILDAKTVWPPINGLTDYHNCQIMADJWICRRECTKVIYVGDNIV
 HSWLIMASVIFHFVACPKFFDPDKTVERAKRGISKIBITNPKEAVKKGADVYS
 NSIVFOAENRHMHQAMVLMHLGK"
 BASE COUNT 1896 a 769 c 878 g 2122 t
 ORIGIN
 Query Match 2.9%; Score 71.8; DB 13; Length 5665;
 Best Local Similarity 44.3%; Pred. No. 0.02; Mismatches 522; Indels 4; Gaps 3;
 Matches 419; Conservative 0;
 DEFINITION Canavalia lineata ornithine carbamoyltransferase OOCt1 (OOCt1)
 ACCESSION AF203688
 VERSION AF203688.1 GI:12003187
 KEYWORDS
 SOURCE
 ORGANISM Canavalia lineata.
 Canavalia lineata: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicots; Rosidae; eurosids I;
 Fabales; Fabaceae; Papilioideae; Canavalia.
 REFERENCE 1. (bases 1 to 5665)
 AUTHORS Yoo,S.K., Lee,Y. and Kwon,Y.M.
 TITLE Properties of ornithine carbamoylases purified from Canavalia
 lineata leaves and their cDNA cloning
 JOURNAL Unpublished
 REFERENCE 2. (bases 1 to 5665)
 AUTHORS Yoo,S.K., Lee,Y. and Kwon,Y.M.
 TITLE Direct Submission
 JOURNAL Submitted (10-NOV-1999) Biology, Seoul National University, San
 56-1, Sinrim-Dong, Kwank-Gu, Seoul 151-742, Korea
 FEATURES
 source
 1. .5665
 /organism="Canavalia lineata"
 /db_xref="txon:28957"
 join((1702..1867,2204..2538,4005..4136,4240..4515,
 5276..5665)
 /gene="OOCt1"
 /product="ornithine carbamoyltransferase OOCt1"
 1702..5665
 /gene="OOCt1"
 join(1714..1867,2204..2538,4005..4136,4240..4515,

Search completed: November 22, 2001, 22:16:10
Job time: 31672 sec

RESULT 1 ATFF9F18

LOCUS ATFF9F18 91740 bp DNA chromosome 4, BAC clone FL9F18 (ESSA)

DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone FL9F18 (ESSA)

ACCESSION AL035605

VERSION AL035605.1

KEYWORDS GI:4468976

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Bukrava; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 91740)

AUTHORS Bevan,M., Koetter,P., Hempel,S., Entian,K.-D., Bancroft,I., Mewes,H.W., Mayer,K.F.X. and Schueler,C.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 91740)

AUTHORS EU Arabidopsis sequencing project.

TITLE JOURNAL Submitted (03-MAR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 1a, D-82152 Martinsried, FRG, E-mail: schueler@mpg.de,mayer@mps.biocbem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Tines Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bsrc.ac.uk

COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosome 4 can be viewed at: <http://webvr.mips.biocbem.mpg.de/proj/thal/>.

FEATURES source

9 76.4 3.1 183250 62 AC012492 Homo sapi

C 10 3.0 879 53 CNS01JRG

C 11 2.9 176174 86 AC007483

C 12 2.9 5665 13 AF03688

C 13 2.9 192929 60 AC005505

C 14 2.9 185800 81 AL512783

C 15 2.9 185800 81 AL512783

C 16 2.8 2305 96 S96842

C 17 2.8 7347 2 AF21124

C 18 2.8 148061 87 AC010727

C 19 2.8 94384 87 AC011718

C 20 2.8 183358 87 AC019041

C 21 2.7 69709 6 CEY6B3B

C 22 2.7 10000 83 CEY37H9_0

C 23 2.7 25516 83 CEY6B3

C 24 2.7 330612 83 CEY8G2

C 25 2.7 160401 87 AC009277

C 26 2.7 253305 96 PRM1ALP7

C 27 2.7 184535 84 CNS05TCJ

C 28 2.7 205546 76 AC083827

C 29 2.7 231912 77 AC087566

C 30 2.7 19517 6 DMU3741

C 31 2.7 150597 84 CNS05TCG

C 32 2.7 184535 84 CNS078R

C 33 2.7 253305 96 PRM1ALP7

C 34 2.7 1141 10 AX083744

C 35 2.7 166666 76 AC073409

C 36 2.6 156060 60 AC004153

C 37 2.6 174384 61 AC009524

C 38 2.6 4601 6 DMU1584

C 39 2.6 136098 86 AC006970

C 40 2.6 6372 6 DDCNPA

C 41 2.6 83440 68 AC024285

C 42 2.6 185603 74 AC073136

C 43 2.6 170157 79 AL355372

C 44 2.6 174565 74 AC069525

C 45 2.6 178087 85 AC005089

ALIGNMENTS

RESULT 1 ATFF9F18

exon 1 /organism="Arabidopsis thaliana"

gene "F19F18.10" /variety="Columbia"

/db_xref="taxon:3702"

/chromosome="4"

/note="Contains Peroxidases signatures [DMIALSGAHTL]

contains EST gb:H76577, Z33618, H37596, R90036, Z35219, AA713320, AA721954"

/db_xref="GI:4468977"

/product="peroxidase, PRX2"

/protein_id="CAB88291.1"

/transl_id="MIVVNTKNTULLLISLICITLDISSLQQLRNPFAGSCIRNEQTVR

NAVYVQQTIPATLRLYHDCFGNGCDASVMTASTNNKAEDHEUNLSLAGD

FQTVIKAKEDVQPKNCVQSCADLTMAADFVNLAGEPOYDVELGDLGJSSTAA

SYGGKLPHPTDDVNUKTLSFLARKGLSANDMALSAGTALGAHCTKVNRYTFNTT

KUDPTVKNDVYVIELKASCSCPRTIDPRVAINMDPTPQRFDNYYTKNLOGQGLFTSDQV

LTDPRKPTDVWANQGQENFAINSMMKLRGVKTSQGNTIRRDCAFN"

exon 2053 : 2268

/gene="F19F18.10"

/number=1

intron 2053 : 2268

/gene="F19F18.10"

/number=1

exon 2269 . 2350

/gene="F19F18.10"

/number=1

exon 2351 . 2551

/gene="F19F18.10"

/number=2

intron 2552 . 2624

/gene="F19F18.10"

/number=2

exon 2625 . 2790

/gene="F19F18.10"

/number=3

intron 2791 . 3002

/gene="F19F18.10"

/number=3

exon 3003 . 3409

/gene="F19F18.10"

/number=4

introns 5135 . 6431

/gene="F19F18.20"

/gene="F19F18.20"

/note="strong similarity to peroxidase, Arabidopsis thaliana, gb:x98314

contains peroxidases signatures [DMIALSGAHTL]"

/codon_start=1

/product="peroxidase-like protein"

/protein_id="CAB38292.1"

/db_xref="GI:4468976"

exon 1 /translation="MVMVMTTNTLLLISLFLAINLSSQLRQDFYACTCPNVEQIVR

NAVKIQQTFTTIPATLRLYHDCFGNGCDASVMTASTNNKAEDHEUNLSLAGD

FDTVKAKEDVQPKNCVQSCADLTMAADFVNLAGEPOYDVELGDLGJSSTAA

SYGTLKPKPFDVNCRKNSCQDILATMATTQDYNLPGQAVELGRGGDLSAS

NVDPTINKVYVTELKASCQDIDRVAINDPQPRODNNVYKLNQGKJLPSDQV

LFTDSRSKPDYDWWANNGLFNQAFISSMKIGGRVGVTGSGNRDRGCAFN"

introns 5135 . 5350

/gene="F19F18.20"

exon 5351 . 5436

/gene="F19F18.20"

/number=1

exon 5437 . 5637

/gene="F19F18.20"

/number=2

5638 . 5726

/gene="F19F18.20"

/number=2

5727 . 5892

/gene="F19F18.20"

/number=2

5727 . 5892

